Lecture 19 Singular Value Decomposition

- Singular value decomposition
- 2 by 2 case
- SVD Theorem
- Similar matrices
- Jordan form

Singular value decomposition

Suppose $A \in \mathbf{R}^{m\times n}$ with $\text{rank}(A) = r$. The singular value decomposition (SVD) of A is to

- choose orthogonal basis $v_1, \cdots, v_r$ of row space of A, and
- choose orthogonal basis $u_1, \cdots, u_r$ of column space of A
- so that $Av_i = \sigma_i u_i$, $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$

In matrix form the equations $Av_i = \sigma_i u_i$ become $AV = U\Sigma$.

The matrix A can be written as

$$A = U\Sigma V^T$$

where $U \in \mathbf{R}^{m\times r}$ and $V \in \mathbf{R}^{n\times r}$ have orthonormal columns.

Application example

Singular value decomposition has many applications in signal processing and control. We consider an example of image compression.

- a (black and white) digital image is a matrix of pixel values
- each pixel contains the grey level
- each picture may have 512 pixels in each row and 256 pixels in each column, a 256 by 512 matrix
- usually in applications large amount of images need to be stored and processed
- compression needed to reduce data to manageable size without losing picture quality

Low rank approximation

Suppose $A \in \mathbf{R}^{256\times 512}$ is a digital image and we have SVD for A as

$$A = U\Sigma V^T.$$

Basic idea:

- $\hat{A}_1 = \sigma_1 u_1 v_1^T$ gives the best rank 1 approximation to A
- compression ratio: $\frac{(256)(512)}{(1+256+512)}$, roughly $170 : 1$
- $\hat{A}_k = \sum_{j=1}^{k} \sigma_j u_j v_j^T$ is the best rank k approximation to A

Let approximation error $E = A - \hat{A}_k$. The approximation is best in the sense that $\sum_i \sum_j |e_{ij}|^2$ is minimized.

2 by 2 case

We will consider $A \in \mathbf{R}^{2\times 2}$ with rank $r = 2$, so A is invertible.

The row space $\mathcal{C}(A^T) = \mathbf{R}^2$ and column space $\mathcal{C}(A) = \mathbf{R}^2$.

We need

- v_1 and v_2 orthonormal
- Av_1 and Av_2 are perpendicular
- $u_1 = Av_1/\|Av_1\|$ and $u_2 = Av_2/\|Av_2\|$

We want to diagonalize A but can not use eigenvectors: A may not be symmetric so eigenvectors are not orthogonal and eigenvalues may not be real.

2 by 2 case

Putting together, with $\|Av_1\| = \sigma_1$ and $\|Av_2\| = \sigma_2$,

$$A \begin{bmatrix} v_1 & v_2 \end{bmatrix} = \begin{bmatrix} \sigma_1 u_1 & \sigma_2 u_2 \end{bmatrix} = \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix}.$$

In matrix form

$$AV = U\Sigma \Leftrightarrow U^{-1}AV = \Sigma \Leftrightarrow U^T AV = \Sigma.$$

- diagonal matrix Σ contains the *singular values* σ_1 and σ_2
- columns of U form an orthogonal basis for $\mathcal{C}(A)$
- columns of V form an orthogonal basis for $\mathcal{C}(A^T)$ $(A^T U = V\Sigma)$

Singular vectors

The singular vectors v_1 and v_2 are the eigenvectors of A^TA with eigenvalues σ_1^2 and σ_2^2:

$$A^TA = (U\Sigma V^T)^T(U\Sigma V^T) = V\Sigma^T\Sigma V^T = V \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix} V^T.$$

The singular vectors u_1 and u_2 are the eigenvectors of AA^T with eigenvalues σ_1^2 and σ_2^2:

$$AA^T = (U\Sigma V^T)(U\Sigma V^T)^T = U\Sigma^T\Sigma U^T = U \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix} U^T.$$

SVD Theorem (2 by 2 case)

Theorem: The singular value decomposition of $A \in \mathbf{R}^{2\times 2}$ with rank $(A) = 2$ has orthogonal matrices U and V so that

$$AV = U\Sigma \iff A = U\Sigma V^{-1} = U\Sigma V^T.$$

where $\Sigma = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix}$ contains the singular values $\sigma_1 > 0$ and $\sigma_2 > 0$.

Example

Find singular value decomposition of $A = \begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix}$.

- $A^T A = \begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix}$
- eigenvalues $\sigma_1^2 = 8$ and $\sigma_2^2 = 2$
- unit eigenvectors $v_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, v_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$
- $Av_1 = \begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 2\sqrt{2} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$
- $Av_2 = \begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \sqrt{2} \begin{bmatrix} 0 \\ 1 \end{bmatrix}$
- $A = U\Sigma V^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2\sqrt{2} & 0 \\ 0 & \sqrt{2} \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$

Interpretation of $A = U\Sigma V^T$

Consider the relation $y = Ax$.

By SVD we decompose the action of A into three simple steps: rotation, scaling and rotation:

- rotate (or reflection) by V^T
- scale along the axes
- rotate by U

The action of A is to transform the unit circle to an ellipse.

Interpretation of $A = U\Sigma V^T$

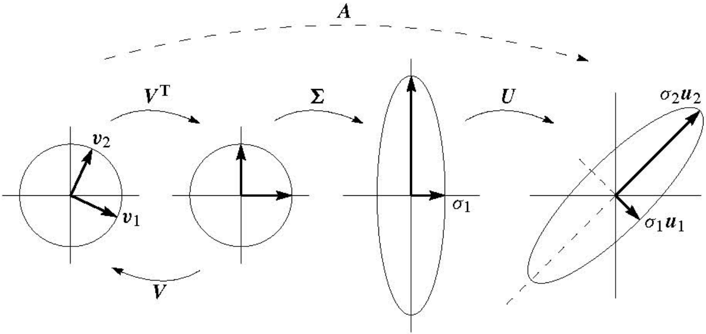

Figure 6.5 U and V are rotations and reflections. Σ is a stretching matrix.

Example

Find singular value decomposition of $A = \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix}$.

- $\text{rank}(A) = 1$
- basis for row space $v_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$
- basis for column space $u_1 = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 \\ 1 \end{bmatrix}$
- $Av_1 = \sqrt{2} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \sigma_1 u_1$, so $\sigma_1 = \sqrt{10}$
- SVD: $A = \sigma_1 u_1 v_1^T$

Example

It is customary to make U and V square. The matrices need a second column.

- v_2 is perpendicular to v_1 so choose $v_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$
- u_2 is perpendicular to u_1 so choose $u_2 = \frac{1}{\sqrt{5}}\begin{bmatrix} 1 \\ -2 \end{bmatrix}$
- v_2 is in $\mathcal{N}(A)$ so $Av_2 = 0$, so $\sigma_2 = 0$
- u_2 is in $\mathcal{N}(A^T)$
- SVD:

$$A = U\Sigma V^T = \frac{1}{\sqrt{5}}\begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix}\begin{bmatrix} \sqrt{10} & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$$

Example

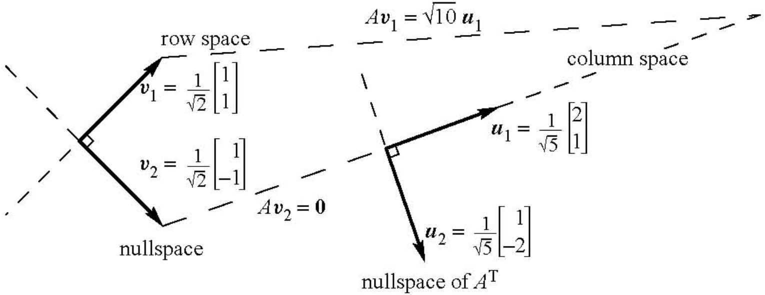

Figure 6.6 The SVD chooses orthonormal bases for 4 subspaces so that $A\boldsymbol{v}_i = \sigma_i \boldsymbol{u}_i$.

SVD Theorem

Theorem: The singular value decomposition of $A \in \mathbf{R}^{m\times n}$, $\text{rank}(A) = r$, has orthogonal matrices U and V so that

$$AV = U\Sigma \;\;\Leftrightarrow\;\; A = U\Sigma V^T = U_1\Sigma_1 V_1^T.$$

- $U = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \in \mathbf{R}^{m\times m}$, $U_1 \in \mathbf{R}^{m\times r}$
- $V = \begin{bmatrix} V_1 & V_2 \end{bmatrix} \in \mathbf{R}^{n\times n}$, $V_1 \in \mathbf{R}^{n\times r}$
- $\Sigma \in \mathbf{R}^{m\times n}$ has the form $\begin{bmatrix} \Sigma_1 & 0_{r\times(n-r)} \\ 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)} \end{bmatrix}$ and

$$\Sigma_1 = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix}, \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0.$$

Orthogonal bases for the 4 spaces

In the SVD $A = U\Sigma V^T$, the orthogonal matrices U and V contain orthonormal bases for the four spaces associated with A.

- columns of V_1 is an orthonormal basis for $\mathcal{C}(A^T)$
- columns of V_2 is an orthonormal basis for $\mathcal{N}(A)$
- columns of U_1 is an orthonormal basis for $\mathcal{C}(A)$
- columns of U_2 is an orthonormal basis for $\mathcal{N}(A^T)$

Proof of SVD Theorem

We have

- rank$(A) =$ rank$(A^T A) = r$
- $A^T A$ has r positive eigenvalues $\sigma_1^2, \cdots, \sigma_r^2$
- singular values $\sigma_1, \cdots, \sigma_r$ are defined

From the equation

$$A^T A v_i = \sigma_i^2 v_i \quad (1)$$

- orthonormal vectors $v_1, \cdots, v_r$ and thus V_1 are defined
- they gives a basis for the row space $\mathcal{C}(A^T)$

Choose orthonormal $v_{r+1}, \cdots, v_n$ as a basis for $\mathcal{N}(A)$.

Thus V_2 and V are defined.

Proof of SVD Theorem

From (1),

- $v_i^T A^T A v_i = \sigma_i^2 v_i^T v_i \Rightarrow \|Av_i\| = \sigma_i$
- $AA^T A v_i = \sigma_i^2 A v_i$
- $u_i = Av_i/\sigma_i, i = 1, \cdots, r$ are orthonormal eigenvectors of AA^T and they form a basis of $\mathcal{C}(A)$, U_1 is defined
- $Av_i = \sigma_i u_i, i = 1, \cdots, r$

Choose orthonormal $u_{r+1}, \cdots, u_m$ as a basis for $\mathcal{N}(A)$, which defines U_2.

Similar matrices

Suppose M is an invertible matrix and $B = M^{-1}AM$.

- we say B is *similar* to A
- if B is similar to A, then A is similar to B
- in differential equations, the expression $M^{-1}AM$ appears when we change variables: consider $\frac{dx}{dt} = Ax$ and let $x = Mz$, then

$$M\frac{dz}{dt} = AMz \Leftrightarrow \frac{dz}{dt} = M^{-1}AMz$$

Invariance of eigenvalues

Fact: Suppose A and B are similar, and $B = M^{-1}AM$. Then (a) A and B have the same eigenvalues and (b) v is an eigenvector of A implies $M^{-1}v$ is an eigenvector of B.

Proof: (a) We have

$$\det(\lambda I - B) = \det(M^{-1}) \det(\lambda I - A) \det(M) = \det(\lambda I - A).$$

(b) Write $A = MBM^{-1}$, then

$$Av = \lambda v \Leftrightarrow MBM^{-1}v = \lambda v \Leftrightarrow B(M^{-1}v) = \lambda(M^{-1}v).$$

This shows that λ is an eigenvalue of B with eigenvector $M^{-1}v$.

Example

Consider the projection matrix $A = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}$.

- the eigenvalues are 1 and 0.
- A is similar to $\Lambda = S^{-1}AS = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$
- choose $M = \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix}$, $M^{-1}AM = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$
- choose $M = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, $M^{-1}AM = \begin{bmatrix} 0.5 & -0.5 \\ -0.5 & 0.5 \end{bmatrix}$
- every 2 by 2 matrix with eigenvalues 1 and 0 is similar to A

Example

The matrix $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ is similar to any nonzero B of the form

$$B = \begin{bmatrix} cd & d^2 \\ -c^2 & -cd \end{bmatrix}.$$

- eigenvalues of A are 0 and 0
- rank$(A) = 1$
- $\det(B) = 0$, rank$(B) = 1$, trace $(B) = 0$.
- B can not be diagonalized
- A is the *Jordan form* of B
- $B = M^{-1}AM$ where $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ with $ad - bc = 1$

Similarity transformation

The formula $B = M^{-1}AM$ is called a *similarity transformation* from A to B.

In the transformation, some things changed and some don't.

Not changed	Changed
eigenvalues	eigenvectors
trace and determinant	nullspace
rank	column space
# of indep. eigenvectors	row space
Jordan form	left nullspace
	singular values

Example

Consider the Jordan matrix $J = \begin{bmatrix} 5 & 1 & 0 \\ 0 & 5 & 1 \\ 0 & 0 & 5 \end{bmatrix}$.

- $J - 5I = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$ has rank 2
- eigenvalues 5, 5, 5 (algebraic multiplicity=3)
- one indep. eigenvector (geometric multiplcity=1)
- $B = M^{-1}JM$ has eigenvalues 5, 5, 5 and $\text{rank}(B - 5I) = 2$.
- $\dim(\mathcal{N}(B - 5I)) = 1$ (one indep. eigenvector)

Example

- J^T is similar to J with $M = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$
- J is similar to every matrix A with eigenvalues 5, 5, 5 and one (independent) eigenvector, *i.e.*, there is an M such that

$$M^{-1}AM = J$$

(this follows from the Jordan Form Theorem)

Example

Consider the differential equation

$$\frac{dx}{dt} = Jx = \begin{bmatrix} 5 & 1 & 0 \\ 0 & 5 & 1 \\ 0 & 0 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \Leftrightarrow \begin{array}{rcl} \frac{dx_1}{dt} & = & 5x_1 + x_2 \\ \frac{dx_2}{dt} & = & 5x_2 + x_3 \\ \frac{dx_3}{dt} & = & 5x_3 \end{array}$$

This is a triangular system and can be solved sequentially from the last equation.

$$\begin{array}{lcl} \frac{dx_3}{dt} = 5x_3 & \Rightarrow & x_3(t) = x_3(0)e^{5t} \\ \frac{dx_2}{dt} = 5x_2 + x_3 & \Rightarrow & x_2(t) = (x_2(0) + tx_3(0))e^{5t} \\ \frac{dx_1}{dt} = 5x_1 + x_2 & \Rightarrow & x_1(t) = (x_1(0) + tx_2(0) + \frac{1}{2}t^2x_3(0))e^{5t} \end{array}$$

Remark: Generalization to Jordan matrix of size n is obvious.

Jordan form

For every $A \in \mathbf{R}^{n \times n}$ we want to choose M so that $M^{-1}AM$ is as *nearly diagonal as possible.*

When A has n independent eigenvectors, $M = S$ and $M^{-1}AM = \Lambda$ is the Jordan form of A

In general, suppose A has s independent eigenvectors.

- A is similar to a matrix with s blocks
- each block is a *Jordan matrix* (called a *Jordan block*): the eigenvalues on the diagonal and the diagonal above it contains 1's
- each block accounts for an eigenvector of A

Jordan Form Theorem

If A has s independent eigenvectors, then it is similar to a matrix J that has s Jordan blocks on its diagonal: There is a matrix M such that

$$M^{-1}AM = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_s \end{bmatrix} = J.$$

Each block in J has one eigenvalue λ_i, one eigenvector, and 1's above the diagonal:

$$\begin{bmatrix} \lambda_i & 1 & & \\ & \cdot & \cdot & \\ & & \cdot & 1 \\ & & & \lambda_i \end{bmatrix}$$

Jordan form

Corollary: If A and B share the same Jordan form, then they are similar

To see this:

$$M_A^{-1} A M_A = J = M_B^{-1} B M_B$$
$$\Rightarrow \; M_B M_A^{-1} A M_A M_B^{-1} = B.$$

Note

- $A^k = M^{-1} J^k M$
- $e^{At} = M^{-1} e^{Jt} M$
- J^k and e^{Jt} are easy to compute

Remark: Numerical computation of M and J is not stable: a slight change in A will separate the repeated eigenvalues.